

Application No. 10/033,586

**Amendments to the Claims:**

This listing of claims will replace all prior versions, and listings, of claims in the application:

**Listing of Claims:**

1. (Currently Amended): A machine-implemented method for extrapolating user profile information from user web page access patterns [[of a user]], comprising:

detecting a set of web pages accessed by a test user having an unknown user profile attribute;

mapping at least a subset of said detected web pages to a first data structure, said first data structure [[identifies]] representing a web page access pattern of said test user;

comparing said first data structure to a plurality of a second data structure to obtain a comparison result, the plurality of said second data structure [[identifies a]] representing clusters of web page access patterns of a sample data set of [[known users, said known]] users having a known user profile attribute in common;

evaluating based on said comparison result the plurality of said second data structure and said first data structure; and

assigning said unknown user profile attribute of said test user from the matching second data structure to said test user in response to said comparison result;

wherein the known user profile attribute in common of the sample data sets corresponds to the unknown user profile attribute of said test user.

2. (Currently Amended): The method of claim 1, wherein said first and second data structures are multi-dimensional vectors, wherein each dimension of said first and said second multi-dimensional vectors corresponds to a different web page,

Application No. 10/033,586

and wherein each dimension of said second data structure corresponds to an average of user multi-dimensional vectors in its corresponding cluster.

3. **(Currently Amended):** The method of claim 2, wherein said comparing step comprises:

determining a distance between said multi-dimensional vectors.

4. **(Currently Amended):** The method of claim 3 [[2]], wherein said [[comparison result is]] determining further comprises computing a cosine of an angle between said multi-dimensional vectors.

5. **(Currently Amended):** The method of claim 2 [[1]], wherein said unknown user profile attribute is demographic information.

6. **(Currently Amended):** The method of claim 5 [[1]], wherein said demographic information [[profile attribute]] is one of [[a]] gender and age [[of said user]].

7. **(Currently Amended):** A machine-implemented method for extrapolating user profile information from user web page access patterns [[of a user]], comprising:

computing bias values for a plurality of web pages;

assigning said bias values to [[a]] the plurality of web pages;

detecting at least a subset of said web pages accessed by a test user having an unknown user profile attribute;

combining said bias values of said subset of web pages to obtain a combination result; and

Application No. 10/033,586

assigning [[a]] a selected user profile attribute to said test user in response to said combination result indicating a positive bias of the certain user profile attribute;

wherein computing said bias values for the plurality of web pages further comprises determining a fraction of users with the selected user profile attribute who visit a selected web page as measured over the plurality of web pages.

8. (Original): The method of claim 7, wherein said combination result is a summation of said bias values of said subset of web pages.

9. (Currently Amended): The method of claim 8 [[7]], wherein said unknown user profile attribute is demographic information.

10. (Currently Amended): The method of claim 9 [[7]], wherein said demographic information is one of gender and age ~~profile attribute is a gender of said user.~~

11. (Currently Amended): A machine-implemented method for extrapolating profile information from user web page access patterns [[of a test user]], comprising:

detecting a set of web pages accessed by a test user having an unknown user profile attribute;

initializing a first set of Expectation Maximization (EM) parameters with data from a training set of users having a known user profile attribute for the unknown user profile attribute of said test user; the data recording web pages accessed by the users in the training set of users;

performing a first EM process using said first set of initialized parameters to obtain a first EM process result that defines a probability of a user profile attribute given said test user; and

Application No. 10/033,586

assigning [[a]] said user profile attribute to the unknown user profile attribute of said test user in response to said first EM process result;

wherein said first EM process is performed while assuming that the unknown user profile attribute of said test user is statistically dependent on whether said test user accesses a web page.

12. (Currently Amended): The method of claim 11, wherein said first EM process comprises the steps of:

performing a first expectation step using said first set of initialized parameters to obtain an expectation result;

performing a first maximization step using said expectation result to obtain a maximization result; and

repeating said expectation and maximization steps, said repeated expectation step uses said maximization result.

13. (Original): The method of claim 12, further comprising:

determining a log-likelihood in place of said repeating step; and

repeating said expectation and maximization steps in response to said log-likelihood.

14. (Original): The method of claim 12, further comprising:

determining an accuracy value on a separate validation data set of said repeating step; and

repeating said expectation and maximization steps in response to said accuracy value.

Application No. 10/033,586

15. (Original): The method of claim 12, wherein said expectation result is a conditional probability of a gender given a web page and a user.

16. (Original): The method of claim 12, wherein said maximization result is a conditional probability of a web page given a gender.

17. (Original): The method of claim 12, wherein said maximization result is a conditional probability of a gender given a user.

18. (Currently Amended): The method of claim 11, wherein:  
~~said first set of EM parameters comprises data from a training set of users, said~~  
 method further comprises ~~[[comprising]]~~:

initializing a second set of EM parameters in place of said assigning step, said second set of EM parameters comprise web page access information for said test user;

performing a second EM process using said first EM process result and said second set of EM parameters to obtain a second EM process result; and

assigning a user profile attribute to the unknown user profile attribute of said test user in response to said second EM process result.

19. (Currently Amended): The method of claim 12 [[18]], wherein ~~at least one profile attribute of said users of said training set is known prior to the performance of said first EM process~~ said first maximization step is given by the following equations:

$$P(s|g) = \frac{\sum_u n(s,u)P(g|s,u)}{\sum_{s'} \sum_u n(s',u)P(g|s',u)}, \text{ and}$$

Application No. 10/033,586

$$P(g|u) = \frac{\sum_s n(s,u)P(g|s,u)}{\sum_{s'} \sum_{g'} n(s',u)P(g'|s',u)}$$

where:

s is a web page

u is a user;

g is gender.

20. (Currently amended): The method of claim 19 [[11]], wherein said first expectation step is given by the following equation:

$$P(g|s,u) = \frac{P(s|g)P(g|u)}{\sum_{g'} (s|g')P(g'|u)}$$

~~said first EM process is performed on a set of users, wherein at least one profile attribute of at least a subset of said set of users is known prior to the performance of said first EM process.~~

21. (Currently Amended): The method of claim 13 [[11]], wherein ~~said first EM process result is a probability of said user profile attribute given said test user~~ said log-likelihood L is determined using the following equation:

$$L = \sum_g \sum_u n(s,u) \log P(s,u)$$

where:

s is a web page

u is a user;

g is gender;

n(s,u) indicates the number of times user u has accessed the web site s.

Application No. 10/033,586

22. **(Currently Amended):** The method of claim 11 ~~[[21]]~~, wherein said user profile attribute is assigned to said test user if said probability is greater than a threshold value.

23. **(Currently Amended):** The method of claim 11, wherein said user profile attribute is demographic information.

24. **(Currently Amended):** The method of claim 23 ~~[[11]]~~, wherein said demographic information is one of gender and age ~~profile attribute is a gender of said test user.~~

25. **(Canceled).**

26. **(Canceled).**

27. **(Canceled).**

28. **(Currently Amended):** An apparatus ~~for extrapolating user profile information from web page access patterns of a user~~, comprising:

a memory, said memory adapted to store program code;

a processor in communication with said memory, said program code capable of programming said processor to perform a method for extrapolating user profile information from user web page access patterns ~~[[of a user]]~~, the method comprising:

detecting a set of web pages accessed by a test user having an unknown user profile attribute;

Application No. 10/033,586

mapping at least a subset of said detected web pages to a first data structure, said first data structure [[identifies]] representing a web page access pattern of said test user;

comparing said first data structure to a plurality of a second data structure to obtain a comparison result, the plurality of said second data structure [[identifies a]] representing clusters of web page access patterns of a set of [[known users, said known]] users having a known user profile attribute in common;

evaluating based on said comparison result the plurality of said second data structure and said first data structure; and

assigning said unknown user profile attribute of said test user from the matching second data structure to said test user in response to said comparison result;

wherein the known user profile attribute in common of the sample data sets corresponds to the unknown user profile attribute of said test user.

29. **(Currently Amended):** An apparatus ~~for extrapolating user profile information from web page access patterns of a user~~, comprising:

a memory, said memory adapted to store program code;

a processor in communication with said memory, said program code capable of programming said processor to perform a method for extrapolating user profile information from user web page access patterns [[of a user]], the method comprising:

detecting a set of web pages accessed by a test user having an unknown user profile attribute;

initializing a first set of Expectation Maximization (EM) parameters with data from a training set of users having a known user profile attribute for the unknown user profile attribute of said test user; the data recording web pages accessed by the users in the training set of users;

performing a first EM process using said first set of initialized parameters to obtain a first EM process result that defines a probability of a user profile attribute given



Application No. 10/033,586

said test user; and

assigning [[a]] said user profile attribute to the unknown user profile attribute of said test user in response to said first EM process result;

wherein said first EM process is performed while assuming that the unknown user profile attribute of said test user is statistically dependent on whether said test user accesses a web page.

30. (Currently Amended): A processor readable storage medium, comprising:

processor readable program code embodied on said processor readable storage medium, said processor readable program code for programming a processor to perform a method for extrapolating user profile information from user web page access patterns [[of a user]], the method comprising:

detecting a set of web pages accessed by a test user having an unknown user profile attribute;

mapping at least a subset of said detected web pages to a first data structure, said first data structure [[identifies]] representing a web page access pattern of said test user;

comparing said first data structure to a plurality of a second data structure to obtain a comparison result, the plurality of said second data structure [[identifies a]] representing clusters of web page access patterns of a sample data set of [[known users, said known]] users having a known user profile attribute in common;

evaluating based on said comparison result the plurality of said second data structure and said first data structure; and

assigning said unknown user profile attribute of said test user from the matching second data structure to said test user in response to said comparison result;

wherein the known user profile attribute in common of the sample data sets corresponds to the unknown user profile attribute of said test user.

Application No. 10/033,586

31. **(Currently Amended):** A processor readable storage medium, comprising:

processor readable program code embodied on said processor readable storage medium, said processor readable program code for programming a processor to perform a method for extrapolating user profile information from user web page access patterns [[of a user]], the method comprising:

detecting a set of web pages accessed by a test user having an unknown user profile attribute;

initializing a first set of Expectation Maximization (EM) parameters with data from a training set of users having a known user profile attribute for the unknown user profile attribute of said test user; the data recording web pages accessed by the users in the training set of users;

performing a first EM process using said first set of initialized parameters to obtain a first EM process result that defines a probability of a user profile attribute given said test user; and

assigning [[a]] said user profile attribute to the unknown user profile attribute of said test user in response to said first EM process result;

wherein said first EM process is performed while assuming that the unknown user profile attribute of said test user is statistically dependent on whether said test user accesses a web page.

Please add the following claims:

32. **(New):** An apparatus, comprising:

a memory, said memory adapted to store program code;

a processor in communication with said memory, said program code capable of programming said processor to perform a method for extrapolating user profile

Application No. 10/033,586

information from user web page access patterns, the method comprising:

computing bias values for a plurality of web pages;

assigning said bias values to the plurality of web pages;

detecting at least a subset of said web pages accessed by a test user having an unknown user profile attribute;

combining said bias values of said subset of web pages to obtain a combination result; and

assigning a selected user profile attribute to said test user in response to said combination result indicating a positive bias of the certain user profile attribute;

wherein computing said bias values for the plurality of web pages further comprises determining a fraction of users with the selected user profile attribute who visit a selected web page as measured over the plurality of web pages.

33. (New): A processor readable storage medium, comprising:

processor readable program code embodied on said processor readable storage medium, said processor readable program code for programming a processor to perform a method for extrapolating user profile information from user web page access patterns, the method comprising:

computing bias values for a plurality of web pages;

assigning said bias values to the plurality of web pages;

detecting at least a subset of said web pages accessed by a test user having an unknown user profile attribute;

combining said bias values of said subset of web pages to obtain a combination result; and

assigning a selected user profile attribute to said test user in response to said combination result indicating a positive bias of the certain user profile attribute;

wherein computing said bias values for the plurality of web pages further

Application No. 10/033,586

comprises determining a fraction of users with the selected user profile attribute who visit a selected web page as measured over the plurality of web pages.

34. (New): The method of claim 7, wherein computing said bias values for the plurality of web pages further comprises computing for each web page the difference between (i) the number of only those users having the selected user profile attribute who visit the selected web page, and (ii) the product of the number of all users who visit the selected web page and the fraction of users with the selected user profile attribute.